

明 細 書

発明の名称： 標的RNAを編集する方法

技術分野

[0001] 本発明は、標的RNAに結合可能なタンパク質を用いるRNA編集技術に関する。本発明は、医療（創薬支援、治療）、農業（農水畜産物生産、育種）、化学（生物学的物質生産）などの幅広い分野で有用である。

背景技術

[0002] 植物のミトコンドリア及び葉緑体ではゲノム中の特定の塩基がRNAレベルで置換されるRNA編集が頻繁に起こり、この現象にRNA結合タンパク質であるペンタトリコペプチドリピート（PPR）タンパク質が介在していることが知られている。

[0003] PPRタンパク質は、このタンパク質を構成するPPRモチーフの構造によって、PとPLSの2つのファミリーに分類される（非特許文献1）。PクラスのPPRタンパク質が標準的な35アミノ酸のPPRモチーフ（P）の単純な繰り返しから構成されている一方、PLSタンパク質は、Pの他に、それに類似したL及びSと呼ばれる2つのモチーフを含んでいる。PLSタンパク質のPPRアレイ（PPRモチーフの並び）は、P1（約35アミノ酸）、L1（約35アミノ酸）、そしてS1（約31アミノ酸）の3つのPPRモチーフがP-L-Sの繰り返し単位として構成され、そのP1L1S1のC末端側に、少し配列の異なるP-L-S、すなわち、P2（35アミノ酸）、L2（36アミノ酸）、そしてS2（32アミノ酸）モチーフが続く。また、P-L-Sの繰り返し単位で構成されているものとは別に、SS(31アミノ酸)が繰り返している場合もある。さらにこの最後のP2L2S2モチーフのC末端側にE1及びE2と呼ばれる2つのPPR様モチーフ、そして136アミノ酸のシチジンデアミナーゼドメイン様配列を持つDYWドメインが続く場合がある（非特許文献2）。

[0004] PPRタンパク質とRNAの相互作用は、各PPRモチーフ中の数カ所のアミノ酸の組み合わせによって結合RNA塩基を指定するPPRコードによって規定されており、これらのアミノ酸が水素結合を介して対応するヌクレオチドと結合する

(特許文献1、特許文献2、非特許文献3、非特許文献4、非特許文献5、非特許文献6、非特許文献7、非特許文献8)。

先行技術文献

特許文献

[0005] 特許文献1：国際公開W02013/058404

特許文献2：国際公開W02014/175284

非特許文献

[0006] 非特許文献1：Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., et al. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16:2089-2103.

非特許文献2：Cheng, S., Gutmann, B., Zhong, X., Ye, Y., Fisher, M. F., Bai, F., Castleden, I., Song, Y., Song, B., Huang, J., et al. (2016). Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J.* 85:532-547.

非特許文献3：Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y. S., Bond, C. S., and Small, I. (2012). A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* 8:e1002910.

非特許文献4：Shen, C., Zhang, D., Guan, Z., Liu, Y., Yang, Z., Yang, Y., Wang, X., Wang, Q., Zhang, Q., Fan, S., et al. (2016). Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. *Nat. Commun.* 7:11285.

非特許文献5：Yagi, Y., Hayashi, S., Kobayashi, K., Hirayama, T., and Nakamura, T. (2013). Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in pl

ants. PLoS One 8:e57286.

非特許文献6 : Kobayashi, T., Yagi, Y., and Nakamura, T. (2019). Comprehensive Prediction of Target RNA Editing Sites for PLS-Class PPR Proteins in *Arabidopsis thaliana*. Plant Cell Physiol. 60:862-874.

非特許文献7 : Yan, J., Yao, Y., Hong, S., Yang, Y., Shen, C., Zhang, Q., Zhang, D., Zou, T., and Yin, P. (2019). Delineation of pentatricopeptide repeat codes for target RNA prediction. Nucleic Acids Res. 47: 3728-3738.

非特許文献8 : Takenaka, M., Zehrmann, A., Brennicke, A., and Graichen, K. (2013). Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. PLoS One 8:e65343.

非特許文献9 : Knie, N., Grewe, F., Fischer, S., and Knoop, V. (2016). Reverse U-to-C editing exceeds C-to-U RNA editing in some ferns – a monilophyte-wide comparison of chloroplast and mitochondrial RNA editing suggests independent evolution of the two processes in both organelles. BMC Evol. Biol. 16:134.

非特許文献10 : Gerke, P., Szovenyi, P., Neubauer, A., Lenz, H., Gutmann, B., McDowell, R., Small, I., Schallenberg-Rudinger, M., and Knoop, V. (2020). Towards a plant model for enigmatic U-to-C RNA editing: the organelle genomes, transcriptomes, editomes and candidate RNA editing factors in the hornwort *Anthoceros agrestis*. New Phytologist 225: 1974-1992.

非特許文献11 : Gutmann B., Royan S., Schallenberg-Rudinger M., Lenz H., Castleden I.R., McDowell R., Vacher M.A., Tonti-Filippini J., Bond C.S., Knoop V., and Small I.D. (2020). The Expansion and Diversification of Pentatricopeptide Repeat RNA-Editing Factors in Plants. Molecular Plant 13:215-230

非特許文献12 : Oldenkott, B., Yang, Y., Lesch, E., Knoop, V., and Scha

llenberg-Rudinger, M. (2019). Plant-type pentatricopeptide repeat proteins with a DYW domain drive C-to-U RNA editing in *Escherichia coli*. *Communications Biology* 2:85.

発明の概要

発明が解決しようとする課題

- [0007] 陸上植物のオルガネラにおいて、RNA塩基がシチジンからウリジンに置換されるC-to-U RNA編集が一般的に起こる。しかし、ツノゴケ類や、小葉類及びシダ植物の一部では、ウリジンがシチジンへ置換されるU-to-C RNA編集も見られる（非特許文献9）。2つの異なるバイオインフォマティクス研究から、U-to-C RNA編集を持つ植物において、標準的なDYWドメインと配列の異なるユニークなDYWドメインが発見された（非特許文献10、非特許文献11）。
- [0008] ヒメツリガネゴケ（*Physcomitrella patens*）由来の2種のDYW：PGタンパク質が、大腸菌内でCからUのRNA編集活性を示すことが報告された（非特許文献12）。陸上植物の進化の初期に分岐した植物群において、DYWドメインは大きく2つのグループに分類される。1つ目はDYW:PG/WWグループと呼ばれ、これには標準的なDYWドメインであるDYW：PG型及びPGボックスにトリプトファン（W）が1つ付加されたDYW:WW型が含まれる。2つ目のグループはDYW：KPと呼ばれ、このDYWドメインはPGボックスとC末端の3つのアミノ酸配列がPG/WWグループと異なる配列を持ち、さらにU-to-C RNA編集を持つ植物のみに存在する。任意のRNA配列内の1塩基を特定の塩基に変える技術は、遺伝子治療や、産業利用における遺伝子変異導入技術において有用である。これまでに、様々なRNA結合分子が開発されているが、DYWドメインと人工的なRNA結合タンパク質を用いることで、任意の標的RNAのシチジン（C）をウリジン（U）に、あるいはその逆方向であるウリジン（U）をシチジン（C）に変換する分子の設計方法は確立されていない。

課題を解決するための手段

- [0009] 本研究において我々は、PPR-DYWタンパク質のC末端DYWドメインを含む部分がRNA編集のモジュラードメイン（modular editing domain）として利用可

能であることを示す。本研究において設計した3種のDYWドメインをそれぞれPR-P又はPLSアレイと融合させた場合、DYW : PG及びDYW : WWドメインはC-to-U、DYW : KPはU-to-CのRNA編集活性を示した。

[0010] 本発明は、以下を提供する。

[1] 下記a、b、c、及びbcのいずれか一のポリペプチドからなるDYWドメインを含む、人工DYWタンパク質を標的RNAに適用する、標的RNAを編集する方法。

a. $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYWを有し、配列番号:1の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

b. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYWを有し、配列番号:2の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

c. $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots x_{c6}x_{c7}x_{c8}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

bc. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots Dx_{bc1}x_{bc2}$ を有し、配列番号:2の90の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

(配列中、xは任意のアミノ酸を表し、…は任意のポリペプチド断片を表す。)

[2] 下記a、b、c、及びbcのいずれか一のポリペプチドからなるDYWドメイン。

a. $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYWを有し、配列番号:1の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

b. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYWを有し、配列番号:2の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

c. $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots x_{c6}x_{c7}x_{c8}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

bc. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots Dx_{bc1}x_{bc2}$ を有し、配列番号:2の90の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

[3] 少なくとも1個のPPRモチーフを含み、標的RNAに配列特異的に結合可能なRNA結合ドメイン、及び2に記載のDYWドメインを含む、DYWタンパク質。

[4] RNA結合ドメインが、PLS型である、2に記載のDYWタンパク質。

[5] 標的RNAを編集する方法であって、

下記c又はbcのポリペプチドからなるDYWドメインを標的RNAに適用して編集標的UをCに変換する工程を含む、方法。

c. $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots x_{c6}x_{c7}x_{c8}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

bc. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots Dx_{bc1}x_{bc2}$ を有し、配列番号:2の90の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

[6] DYWドメインが、少なくとも1個のPPRモチーフを含み、PPR-codeのルールにしたがって標的RNAに配列特異的に結合可能なRNA結合ドメインに融合されている、5に記載の方法。

[7] 真核細胞においてRNAを編集するための、2に記載のDYWドメインを含む、組成物。

[8] 2に記載のDYWドメイン、又は3又は4に記載のDYWタンパク質をコードする核酸。

[9] 8に記載の核酸を含む、ベクター。

[10] 9に記載のベクターを含む、細胞（ヒト個体は除く。）。

[0011] [1] 下記a～cのいずれか一のポリペプチドからなるDYWドメインを含む、人

工DYWタンパク質を標的RNAに適用する、標的RNAを編集する方法。

a. $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots HSE \cdots Cx_{a17}x_{a18}CH \cdots$ DYWを有し、配列番号:1の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

b. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots HSE \cdots Cx_{b17}x_{b18}CH \cdots$ DYWを有し、配列番号:2の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

c. $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots HAE \cdots Cx_{c4}x_{c5}CH \cdots Dx_{c6}x_{c7}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

(配列中、xは任意のアミノ酸を表し、 $\cdots$ は任意のポリペプチド断片を表す。)

[2] 下記a~cのいずれか一のポリペプチドからなるDYWドメイン。

a. $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots HSE \cdots Cx_{a17}x_{a18}CH \cdots$ DYWを有し、配列番号:1の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

b. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots HSE \cdots Cx_{b17}x_{b18}CH \cdots$ DYWを有し、配列番号:2の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

c. $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots HAE \cdots Cx_{c4}x_{c5}CH \cdots Dx_{c6}x_{c7}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

[3] 少なくとも1個のPPRモチーフを含み、標的RNAに配列特異的に結合可能なRNA結合ドメイン、及び2に記載のDYWドメインを含む、DYWタンパク質。

[4] RNA結合ドメインが、PLS型である、2に記載のDYWタンパク質。

[5] 標的RNAを編集する方法であって、

下記cのポリペプチドからなるDYWドメインを標的RNAに適用して編集標的UをCに変換する工程を含む、方法。

c. $KPAx_{c1}Ax_{c2}IEx_{c3}\cdots HAE \cdots Cx_{c4}x_{c5}CH \cdots Dx_{c6}x_{c7}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

[6] DYWドメインが、少なくとも1個のPPRモチーフを含み、PPR-codeのルールにしたがって標的RNAに配列特異的に結合可能なRNA結合ドメインに融合されている、5に記載の方法。

[7] 真核細胞においてRNAを編集するための、2に記載のDYWドメインを含む、組成物。

[8] 2に記載のDYWドメイン、又は3又は4に記載のDYWタンパク質をコードする核酸。

[9] 8に記載の核酸を含む、ベクター。

[10] 9に記載のベクターを含む、細胞（ヒト個体は除く。）。

発明の効果

[0012] 本発明により、標的RNAに含まれる編集標的CをUに、又は編集標的UをCに変換できる。

図面の簡単な説明

[0013] [図1]DYWタンパク質におけるC末端ドメインの近似的最尤系統樹。(a) DYW : PG型、(b) DYW : WW型及び(c) DYW : KP型ドメインの系統樹をFastTreeを用いて作成した。DYW : PGドメインは、小葉類由来のDYW : PGタンパク質に基づいて設計した。DYW : WW及びDYW : KPドメインを設計するために選んだタンパク質の系統群（クレード）を黒色の線で示している。系統樹はiTOL (Letunic, I. and Bork, P. (2016) Nucleic Acids Res. 44 W242-245) を使って可視化しており、赤はツノゴケ類、緑は小葉類、青はシダ植物由来のタンパク質を示している。

[図2]本研究で設計したPPRタンパク質の構成。(a) PPRタンパク質は、N末端側にチオレドキシン、His-タグ及びTEV部位、それに続いて、P又はPLSモチーフからなるPPRアレイ、最後に、DYWドメイン（DYW : PG、DYW : WW又はDYW : KP）を含む。PPRタンパク質の標的配列（RNA編集部位を含む）は終止コドンの

下流に挿入した。(b) P、P1、L1、S1及びP2モチーフ内のRNA認識に関わる4番目とii番目のアミノ酸。L2、S2、E1及びE2の4番目とii番目のアミノ酸はCLB19のものと同一である (Chateigner-Boutin, A.L., et al. (2008). Plant J. 56 590-602.)。

[図3]大腸菌を使ったRNA編集アッセイ。(a) 各PPR-DYWのDNA上における標的塩基 (C又はT)。(b) PPR-DYW:PG及びWWのC-to-U RNA編集活性と (c) PPR-DYW:KPのU-to-C RNA編集活性を示す。3回独立して行った実験結果のうち1つを例として示している。(d) C-to-U及び (e) U-to-C RNA編集効率を3回測定した平均値 (ただし、P-DYW:KPは2回)。エラーバーは標準偏差を示す。

[図4-1]選抜した各DYWドメイン系統群におけるアミノ酸の出現頻度。WebLogoにより作成したシーケンスロゴで視覚化した。図4-2及び図4-3において同じ。

(a) DYW:PG

[図4-2] (b) DYW:WW

[図4-3] (c) DYW:KP

[図5]HEK293T細胞におけるRNA編集活性(C to U またはU to C)。PLS型にPG1もしくはWW1ドメイン(a)、またはKP1ドメイン(b)を融合したときの標的部位のシーケンシング結果。3回の独立した実験におけるRNA編集活性 (c)。棒グラフは平均を示し、各点は3回の実験におけるRNA編集活性を示す。左の棒グラフは、C to U編集活性、右の棒グラフはU to C編集活性を示す。

[図6]ドメインスワッピングによるKPドメイン性能の向上。KP1のActive site (HxEx_nCxxCH) を含む中央部分を残し、前後のドメインについてWW1ドメインとスワップしたChimeric KP1aのRNA編集活性について調べた。ドメインスワッピングの模式図 (a)、実際のRNA編集活性 (b, c)。

[図7]KPドメイン変異体のRNA編集活性測定。KP1, KP2, KP3, KP4の大腸菌でのRNA編集活性(a)とHEK293TでのRNA編集活性(b)、KP5～KP23のHEK293TでのRNA編集活性(c)。

[図8]PGドメイン変異体のRNA編集活性測定。PG1, PG2, の大腸菌でのRNA編集活性(a)とHEK293TでのRNA編集活性(b)。PG3～PG13のHEK293TでのRNA編集活性

[図9]WWドメイン変異体のRNA編集活性測定。WW1, WW2, の大腸菌でのRNA編集活性(a)とHEK293TでのRNA編集活性(b)。WW3~WW14のHEK293TでのRNA編集活性
 [図10]ヒトミトコンドリアRNA編集。標的配列情報(a)。RNA編集活性(b, c)
発明を実施するための形態

[0014] 本発明は、下記a、b、c、及びbcのいずれか一のポリペプチドからなるDYWドメインを含むDYWタンパク質を標的RNAに適用する、標的RNAを編集する方法に関する。

a. $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYWを有し、配列番号:1の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

b. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYWを有し、配列番号:2の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

c. $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots x_{c6}x_{c7}x_{c8}$ を有し、配列番号:3の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

bc. $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots Dx_{bc1}x_{bc2}$ を有し、配列番号:2の90の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチド

[0015] 本発明に関し、C-to-U/U-to-C編集活性とは、対象ポリペプチドを、標的RNAに配列特異的に結合可能なRNA結合ドメインのC末端側に連結して編集アッセイを行ったとき、標的RNAに含まれる編集標的CをUに、又は編集標的UをCに変換できる活性をいう。変換は、適切な条件において編集標的塩基の少なくとも約3%、好ましくは約5%が目的の塩基に置き換わっていればよい。

[0016] [DYWドメイン]

$x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYW、 $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYW、及び $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots x_{c6}x_{c7}x_{c8}$ は、いずれもアミノ酸配列を表す。配列中、xは、それぞれ独立に、任

意のアミノ酸を表し、…は、それぞれ独立に、任意の長さの任意のアミノ酸配列からなるポリペプチド断片を表す。本発明に関し、DYWドメインは、これらの3つのアミノ酸配列のいずれか一つで表すことができる。特に、 $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYWからなるDYWドメインをDYW:PG、 $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYWからなるDYWドメインをDYW:WW、及び $KPAx_{c1}Ax_{c2}IEx_{c3} \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots x_{c6}x_{c7}x_{c8}$ からなるDYWドメインをDYW:KPと表すことがある。

[0017] DYWドメインは、N末端の約15アミノ酸からなるPGボックスを含む領域、中心部のzinc結合ドメイン ($HxEx_nCxxCH$ 、 x_n は任意のアミノ酸の任意の数 n 個の連なりである。)、及びC末端のDYWの3つの領域を有する。zinc結合ドメインはさらに、HxE領域とCxxCH領域に分けることができる。各DYWドメインのこれらの領域は、下表のように表すことができる。

[0018] [表1]

	PGボックスを含む領域	HxE領域	CxxCH領域	DYW
DYW:PG	$x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP$	$Hx_{aa}E$	$Cx_{a17}x_{a18}CH$	DYW
DYW:WW	$x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP$	$Hx_{bb}E$	$Cx_{b17}x_{b18}CH$	DYW
DYW:KP	$KPAx_{c1}Ax_{c2}IEx_{c3}$	$Hx_{cc}E$	$Cx_{c4}x_{c5}CH$	$x_{c6}x_{c7}x_{c8}$

[0019] (DYW:PG)

DYW:PGは、 $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYWからなるポリペプチドである。好ましくは、 $x_{a1}PGx_{a2}SWIEx_{a3}-x_{a16}HP \cdots Hx_{aa}E \cdots Cx_{a17}x_{a18}CH \cdots$ DYWを有し、配列番号:1の配列と配列同一性（〔用語〕の項で詳述する）を有し、かつC-to-U/U-to-C編集活性を有するポリペプチドである。DYW:PGは、編集標的CをUに変換する活性（C-to-U編集活性）を有する。配列番号:1には、本明細書の実施例の項に示した実験で使用した、全長136アミノ酸長からなるDYW:PGの配列を示した。この配列は、本願が初めて開示するものであり、新規である。

[0020] DYW:PGの全長は、C-to-U編集活性を発揮できる限り特に限定されないが、例えば110～160アミノ酸長であり、好ましくは124～148アミノ酸長であり、

より好ましくは128～144アミノ酸長であり、さらに好ましくは132～140アミノ酸長である。

[0021] DYW:PGのPGボックスを含む領域 (x_{a1} PG x_{a2} SWIE x_{a3} ～ x_{a16} HP) において：

x_{a1} は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはE（グルタミン酸）又はそれに性質の似たアミノ酸であり、より好ましくはGである。

x_{a2} は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはC（システイン）又はそれに性質の似たアミノ酸であり、より好ましくはCである。

x_{a3} ～ x_{a16} の各々のアミノ酸は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:1の配列の位置9～22の対応するアミノ酸と同じか、又は対応するアミノ酸と性質の似たアミノ酸であり、より好ましくは配列番号:1の配列の位置9～22の対応するアミノ酸と同じアミノ酸である。

[0022] 好ましい態様の一つにおいて、DYW:PGのHxE領域は、他の領域がどのような場合であってもHSEである。

[0023] DYW:PGのCxxCH領域、すなわちC x_{a17} x_{a18} CHにおいて：

x_{a17} は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはG（グリシン）又はそれに性質の似たアミノ酸であり、より好ましくはGである。

x_{a18} は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはD（アスパラギン酸）又はそれに性質の似たアミノ酸であり、より好ましくはDである。

[0024] DYW:PGにおいて、PGボックスを含む領域とH x_{aa} Eを結合する…部分、H x_{aa} E領域とCxxCH領域を結合する…部分、及びCxxCH領域とDYWを連結する…部分を、順に、第一の連結部、第二の連結部、及び第三の連結部という（他のDYWドメインにおいても同じ。）。

[0025] DYW:PGの第一の連結部の全長は、DYW:PGとしてC-to-U編集活性を発揮でき

る限り特に限定されないが、例えば39～47アミノ酸長であり、好ましくは40～46アミノ酸長であり、より好ましくは41～45アミノ酸長であり、さらに好ましくは42～44アミノ酸長である。第一の連結部のアミノ酸配列は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:1の配列の位置25～67の部分と同じか、その部分配列において1～22個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0026] DYW:PGの第一の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、43アミノ酸長の下式で表されるポリペプチドである。

[0027] $N_{a25}-N_{a26}-N_{a27}-\cdots-N_{a65}-N_{a66}-N_{a67}$

[0028] 上記ポリペプチドは、好ましくは配列番号:1の配列の位置25～67の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:PGとしてC-to-U編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{a29} 、 N_{a30} 、 N_{a32} 、 N_{a33} 、 N_{a35} 、 N_{a36} 、 N_{a40} 、 N_{a44} 、 N_{a45} 、 N_{a47} 、 N_{a48} 、 N_{a52} 、 N_{a53} 、 N_{a54} 、 N_{a55} 、 N_{a58} 、 N_{a61} 、 N_{a65} 、 N_{a67} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0029] DYW:PGの第二の連結部の全長は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、例えば21～29アミノ酸長であり、好ましくは22～28アミノ酸長であり、より好ましくは23～27アミノ酸長であり、さらに好ましくは24～26アミノ酸長である。第二の連結部のアミノ酸配列は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:1の配列の位置71～95の部分と同じか、その部分配列において1～13個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0030] DYW:PGの第二の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、25アミノ酸長の下式で表されるポリペプチドである。

[0031] $N_{a71}-N_{a72}-N_{a73}-\cdots-N_{a93}-N_{a94}-N_{a95}$

[0032] 上記ポリペプチドは、好ましくは配列番号:1の配列の位置71～95の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:PGとしてC-to-U編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{a71} 、 N_{a72} 、 N_{a73} 、 N_{a76} 、 N_{a77} 、 N_{a78} 、 N_{a79} 、 N_{a81} 、 N_{a82} 、 N_{a86} 、 N_{a88} 、 N_{a89} 、 N_{a91} 、 N_{a92} 、 N_{a93} 、 N_{a94} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0033] DYW:PGの第三の連結部の全長は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、例えば29～37アミノ酸長であり、好ましくは30～36アミノ酸長であり、より好ましくは31～35アミノ酸長であり、さらに好ましくは32～34アミノ酸長である。第三の連結部のアミノ酸配列は、DYW:PGとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:1の配列の位置101～133の部分と同じか、その部分配列において1～17個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0034] DYW:PGの第三の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、33アミノ酸長の下式で表されるポリペプチドである。

[0035] $N_{a101}-N_{a102}-N_{a103}-\cdots-N_{a131}-N_{a132}-N_{a133}$

[0036] 上記ポリペプチドは、好ましくは配列番号:1の配列の位置101～133の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:PGとしてC-to-U編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{a102} 、 N_{a103} 、 N_{a104} 、 N_{a105} 、 N_{a106} 、 N_{a107} 、 N_{a108} 、 N_{a109} 、 N_{a110} 、 N_{a111} 、 N_{a112} 、 N_{a113} 、 N_{a114} 、 N_{a115} 、 N_{a116} 、 N_{a117} 、 N_{a118} 、 N_{a119} 、 N_{a120} 、 N_{a121} 、 N_{a122} 、 N_{a123} 、 N_{a124} 、 N_{a125} 、 N_{a126} 、 N_{a127} 、 N_{a128} 、 N_{a129} 、 N_{a130} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

$_{104}$ 、 N_{a107} 、 N_{a112} 、 N_{a114} 、 N_{a117} 、 N_{a118} 、 N_{a121} 、 N_{a122} 、 N_{a123} 、 N_{a124} 、 N_{a125} 、 N_{a128} 、 N_{a130} 、 N_{a131} 、 N_{a132})は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0037] 配列番号:1の配列からなるPGドメインに変異を導入することにより、RNA編集活性を向上することができる。このような変異を導入したドメインの好ましい例は、本明細書の実施例の項に示したPG11（配列番号:50のアミノ酸配列からなるポリペプチド）である。

[0038] (DYW:WW)

DYW:WWは、 $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYWからなるポリペプチドである。好ましくは、 $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{bb}E \cdots Cx_{b17}x_{b18}CH \cdots$ DYWを有し、配列番号:2の配列と配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチドである。DYW:WWは、編集標的CをUに変換する活性（C-to-U編集活性）を有する。配列番号:2には、本明細書の実施例の項に示した実験で使用した、全長137アミノ酸長からなるDYW:WWの配列を示した。この配列は、本願が初めて開示するものであり、新規である。

[0039] DYW:WWの全長は、C-to-U編集活性を発揮できる限り特に限定されないが、例えば110～160アミノ酸長であり、好ましくは125～149アミノ酸長であり、より好ましくは129～145アミノ酸長であり、さらに好ましくは133～141アミノ酸長である。

[0040] DYW:WWのPGボックスを含む領域、すなわち $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP$ において、WTDからなる部分は、WSDであってもよい。

[0041] DYW:WWのPGボックスを含む領域において：

x_{b1} は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはK（リシン）又はそれに性質の似たアミノ酸であり、より好ましくはKである。

x_{b2} は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはQ（グルタミン）又はそれに性質の似たアミノ酸であり、より好ましくはQである。

x_{b3} - x_{b16} の各々のアミノ酸は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:2の配列の位置10~23の対応するアミノ酸と同じか、又は対応するアミノ酸と性質の似たアミノ酸であり、より好ましくは配列番号:2の配列の位置10~23の対応するアミノ酸と同じアミノ酸である。

[0042] 好ましい態様の一つにおいて、DYW:WWのHxE領域は、他の部分の配列がどのような場合であっても、HSEである。

[0043] DYW:WWのCxxCH領域、すなわちC x_{b17} x_{b18} CHにおいて：

x_{b17} は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはD（アスパラギン酸）又はそれに性質の似たアミノ酸であり、より好ましくはDである。

x_{b18} は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくはD又はそれに性質の似たアミノ酸であり、より好ましくはDである。

[0044] DYW:WWの第一の連結部の全長は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、例えば39~47アミノ酸長であり、好ましくは40~46アミノ酸長であり、より好ましくは41~45アミノ酸長であり、さらに好ましくは42~44アミノ酸長である。第一の連結部のアミノ酸配列は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:2の配列の位置25~67の部分と同じか、その部分配列において1~22個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0045] DYW:WWの第一の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、43アミノ酸長の下式で表されるポリペプチドである。

[0046] $N_{b26}-N_{b27}-N_{b28}-\cdots-N_{b66}-N_{b67}-N_{b68}$

[0047] 上記ポリペプチドは、好ましくは配列番号:2の配列の位置26~68の部分と

同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:PGとしてC-to-U編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{b26} 、 N_{b30} 、 N_{b33} 、 N_{b34} 、 N_{b37} 、 N_{b41} 、 N_{b45} 、 N_{b46} 、 N_{b48} 、 N_{b49} 、 N_{b51} 、 N_{b52} 、 N_{b53} 、 N_{b55} 、 N_{b56} 、 N_{b57} 、 N_{b59} 、 N_{b61} 、 N_{b62} 、 N_{b63} 、 N_{b64} 、 N_{b66} 、 N_{b67} 、 N_{b68} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0048] DYW:WWの第二の連結部の全長は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、例えば21～29アミノ酸長であり、好ましくは22～28アミノ酸長であり、より好ましくは23～27アミノ酸長であり、さらに好ましくは24～26アミノ酸長である。第二の連結部のアミノ酸配列は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:2の配列の位置71～95の部分と同じか、その部分配列において1～13個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0049] DYW:WWの第二の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、25アミノ酸長の下式で表されるポリペプチドである。

[0050] $N_{b72}-N_{b73}-N_{b74}-\cdots-N_{b94}-N_{b95}-N_{b96}$

[0051] 上記ポリペプチドは、好ましくは配列番号:2の配列の位置72～96の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:WWとしてC-to-U編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{b72} 、 N_{b73} 、 N_{b74} 、 N_{b75} 、 N_{b77} 、 N_{b78} 、 N_{b79} 、 N_{b81} 、 N_{b82} 、 N_{b84} 、 N_{b88} 、 N_{b89} 、 N_{b90} 、 N_{b91} 、 N_{b92} 、 N_{b93} 、 N_{b94} 、 N_{b95} 、 N_{b96} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0052] DYW:WWの第三の連結部の全長は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、例えば29～37アミノ酸長であり、好ましくは30

～36アミノ酸長であり、より好ましくは31～35アミノ酸長であり、さらに好ましくは32～34アミノ酸長である。第三の連結部のアミノ酸配列は、DYW:WWとしてC-to-U編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:2の配列の位置101～133の部分と同じか、その部分配列において1～17個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0053] DYW:WWの第三の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、33アミノ酸長の下式で表されるポリペプチドである。

[0054] $N_{b102}-N_{b103}-N_{b104}-\cdots-N_{b132}-N_{b133}-N_{b134}$

[0055] 上記ポリペプチドは、好ましくは配列番号:2の配列の位置102～134の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:WWとしてC-to-U編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{b104} 、 N_{b105} 、 N_{b107} 、 N_{b108} 、 N_{b109} 、 N_{b110} 、 N_{b111} 、 N_{b113} 、 N_{b115} 、 N_{b116} 、 N_{b117} 、 N_{b118} 、 N_{b119} 、 N_{b121} 、 N_{b122} 、 N_{b123} 、 N_{b124} 、 N_{b126} 、 N_{b129} 、 N_{b131} 、 N_{b132} 、 N_{b133} 、 N_{b134} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0056] 配列番号:2の配列からなるWWドメインに変異を導入することにより、RNA編集活性を向上することができる。このような変異を導入したドメインの好ましい例は、本明細書の実施例の項に示したWW2～11及びWW13であり、特に編集活性が高いのはWW11（配列番号:63のアミノ酸配列からなるポリペプチド）である。

[0057] (DYW:KP)

DYW:KPは、 $KPAx_{c1}Ax_{c2}IEx_{c3}\cdots Hx_{cc}E\cdots Cx_{c4}x_{c5}CH\cdots x_{c6}x_{c7}x_{c8}$ からなるポリペプチドである。好ましくは、 $KPAx_{c1}Ax_{c2}IEx_{c3}\cdots Hx_{cc}E\cdots Cx_{c4}x_{c5}CH\cdots x_{c6}x_{c7}x_{c8}$ を有し、配列番号:3の配列と配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチドである。DYW:KPは、編集標的UをCに変換する活性（U-t

o-C編集活性)を有する。配列番号:3には、本明細書の実施例の項に示した実験で使用した、全長133アミノ酸長からなるDYW:KPの配列を示した。この配列は、本願が初めて開示するものであり、新規である。

[0058] DYW:KPの全長は、U-to-C編集活性を発揮できる限り特に限定されないが、例えば110~160アミノ酸長であり、好ましくは121~145アミノ酸長であり、より好ましくは125~141アミノ酸長であり、さらに好ましくは129~137アミノ酸長である。

[0059] DYW:KPのPGボックスを含む領域、すなわち $KPAx_{c1}Ax_{c2}IEx_{c3}$ において：

x_{c1} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはS（セリン）又はそれに性質の似たアミノ酸であり、より好ましくはSである。

x_{c2} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはL（ロイシン）又はそれに性質の似たアミノ酸であり、より好ましくはLである。

x_{c3} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはV（バリン）又はそれに性質の似たアミノ酸であり、より好ましくはVである。

[0060] 好ましい態様の一つにおいて、DYW:KPのHxE領域は、他の部分の配列がどのような場合であっても、HAEである。

[0061] DYW:KPのCxxCH領域、すなわち $Cx_{c4}x_{c5}CH$ において：

x_{c4} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはN（アスパラギン）又はそれに性質の似たアミノ酸であり、より好ましくはNである。

x_{c5} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはD（アスパラギン酸）又はそれに性質の似たアミノ酸であり、より好ましくはDである。

[0062] DYW:KPのDYWに相当する部分、すなわち $x_{c6}x_{c7}x_{c8}$ において：

x_{c6} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが

、好ましくはD（アスパラギン酸）又はそれに性質の似たアミノ酸であり、より好ましくはDである。

x_{c7} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはM（メチオニン）又はそれに性質の似たアミノ酸であり、より好ましくはMである。

x_{b8} は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくはF（フェニルアラニン）又はそれに性質の似たアミノ酸であり、より好ましくはFである。

好ましい態様の一つにおいて、 $x_{c6}x_{c7}x_{c8}$ は、他の部分の配列がどのような場合であっても、 $Dx_{c7}x_{c8}$ である。

別の好ましい態様の一つにおいて、 $x_{c6}x_{c7}x_{c8}$ は、他の部分の配列がどのような場合であっても、GRPである。

[0063] DYW:KPの第一の連結部の全長は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、例えば51～59アミノ酸長であり、好ましくは52～58アミノ酸長であり、より好ましくは53～57アミノ酸長であり、さらに好ましくは54～56アミノ酸長である。第一の連結部のアミノ酸配列は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:3の配列の位置10～64の部分と同じか、その部分配列において1～28個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0064] DYW:KPの第一の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、55アミノ酸長の下式で表されるポリペプチドである。

[0065] $N_{c10}-N_{c11}-N_{c12}-\cdots-N_{c62}-N_{c63}-N_{c64}$

[0066] 上記ポリペプチドは、好ましくは配列番号:3の配列の位置10～64の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:KPとしてU-to-C編集活性を発揮できるものである。このときアミノ酸の置

換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{c10} 、 N_{c13} 、 N_{c14} 、 N_{c15} 、 N_{c16} 、 N_{c17} 、 N_{c18} 、 N_{c19} 、 N_{c25} 、 N_{c26} 、 N_{c29} 、 N_{c30} 、 N_{c33} 、 N_{c34} 、 N_{c36} 、 N_{c38} 、 N_{c41} 、 N_{c42} 、 N_{c44} 、 N_{c45} 、 N_{c47} 、 N_{c49} 、 N_{c55} 、 N_{c58} 、 N_{c59} 、 N_{c62} 、 N_{c63} 、 N_{c64} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0067] DYW:KPの第二の連結部の全長は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、例えば21～29アミノ酸長であり、好ましくは22～28アミノ酸長であり、より好ましくは23～27アミノ酸長であり、さらに好ましくは24～26アミノ酸長である。第二の連結部のアミノ酸配列は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:3の配列の位置68～92の部分と同じか、その部分配列において1～13個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0068] DYW:KPの第二の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、25アミノ酸長の下式で表されるポリペプチドである。

[0069] $N_{c68}-N_{c69}-N_{c70}-\cdots-N_{c90}-N_{c91}-N_{c92}$

[0070] 上記ポリペプチドは、好ましくは配列番号:3の配列の位置68～92の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:KPとしてU-to-C編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{c68} 、 N_{c70} 、 N_{c71} 、 N_{c72} 、 N_{c73} 、 N_{c74} 、 N_{c75} 、 N_{c76} 、 N_{c77} 、 N_{c78} 、 N_{c79} 、 N_{c80} 、 N_{c81} 、 N_{c83} 、 N_{c84} 、 N_{c85} 、 N_{c86} 、 N_{c87} 、 N_{c88} 、 N_{c89} 、 N_{c90} 、 N_{c91} 、 N_{c92} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0071] DYW:KPの第三の連結部の全長は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、例えば29～37アミノ酸長であり、好ましくは30～36アミノ酸長であり、より好ましくは31～35アミノ酸長であり、さらに好

ましくは32～34アミノ酸長である。第三の連結部のアミノ酸配列は、DYW:KPとしてU-to-C編集活性を発揮できる限り特に限定されないが、好ましくは配列番号:3の配列の位置98～130の部分と同じか、その部分配列において1～17個のアミノ酸を置換、欠失、又は付加した配列か、又はその部分配列と配列同一性を有する配列であり、より好ましくは、その部分配列と同じ配列である。

[0072] DYW:KPの第三の連結部の好ましい態様の一つは、DYWドメインの他の部分の配列がどのような場合であっても、33アミノ酸長の下式で表されるポリペプチドである。

[0073] $N_{c98}-N_{c99}-N_{c100}-\cdots-N_{c128}-N_{c129}-N_{c130}$

[0074] 上記ポリペプチドは、好ましくは配列番号:3の配列の位置98～130の部分と同じか、その部分配列において複数個のアミノ酸を置換した配列であってDYW:KPとしてU-to-C編集活性を発揮できるものである。このときアミノ酸の置換は、図4の対応する位置においてbits値が大きいアミノ酸（例えば、 N_{c98} 、 N_{c100} 、 N_{c101} 、 N_{c103} 、 N_{c104} 、 N_{c105} 、 N_{c107} 、 N_{c109} 、 N_{c110} 、 N_{c111} 、 N_{c112} 、 N_{c114} 、 N_{c115} 、 N_{c117} 、 N_{c118} 、 N_{c119} 、 N_{c120} 、 N_{c121} 、 N_{c122} 、 N_{c123} 、 N_{c124} 、 N_{c125} 、 N_{c127} 、 N_{c129} 、 N_{c130} ）は図4と同じであり、それ以外のアミノ酸が置換されるように行われていることが好ましい。

[0075] 配列番号:3の配列からなるKPドメインに変異を導入することにより、編集活性を向上することができる。このような変異を導入したドメインの好ましい例は、本明細書の実施例の項に示したKP2～23（配列番号:68～89）である。KP22（配列番号:88）が最もU to Cへの編集活性が高く、かつC to Uへの編集活性が低く、配列番号:3の配列からなるKPドメインと比較してRNA編集活性が改善されている。

[0076] (Chimeric DYW)

DYWドメインは、アミノ酸配列の保存性からいくつかの領域に区分けすることができる。これらの各領域を交換したChimeric DYWにより、C-to-U編集活性又はU-to-C編集活性を向上しうる。好ましい態様の一つは、DYW:WWのPGボ

ックスを含む領域、すなわち $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP$ とDYWを、DYW:KPのその他の領域($\cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots$)に融合したものである。このときの x_{b1} 、 x_{b2} 、 $x_{b3}-x_{b16}$ は、DYW:WWについての上記の説明のとおりである。また第一の連結部、第二の連結部、第三の連結部は、DYW:KPについての上記の説明のとおりである。DYW部分は、 $Dx_{bc1}x_{bc2}$ であってもよい。全長は、U-to-C編集活性を発揮できる限り特に限定されないが、例えば例えば110~160アミノ酸長であり、好ましくは125~149アミノ酸長であり、より好ましくは129~145アミノ酸長であり、さらに好ましくは133~141アミノ酸長である。

[0077] 好ましいChimericドメインの一つは、 $x_{b1}PGx_{b2}SWWTDx_{b3}-x_{b16}HP \cdots Hx_{cc}E \cdots Cx_{c4}x_{c5}CH \cdots Dx_{bc1}x_{bc2}$ を有し、配列番号:2の90の配列と少なくとも40%の配列同一性を有し、かつC-to-U/U-to-C編集活性を有するポリペプチドからなる。

[0078] 特に好ましいChimericドメインの一つは、配列番号:90の配列と配列同一性を有し、かつU-to-C編集活性を有するポリペプチドである。配列番号:90には、本明細書の実施例の項に示した実験で使したChimericドメインの配列を示した。このドメインは、配列番号:3の配列からなるDYW:KPよりも高いU to C編集活性が認められ、さらにC to U編集活性がほとんどない。

[0079] (公知の配列との比較)

前述のとおり、配列番号:1、2、3の配列からなるDYWドメインは新規なものである。以下に、PPRデータベース (<https://ppr.plantenergy.uwa.edu.au/onekp/>; 前掲非特許文献11) を用いて調査した結果を示す。

[0080]

[表2-1]

DYW:PG(配列番号:1)

配列名称	由来	同一性
- PYHZ-2068577-R	Isoetes, 小葉類	78%
- ENQF-2003735-R	Lycopodium annotinum, 小葉類	

DYW:WW(配列番号:2)

- ANON-2000164-R	Leiosporoceros dussii, ツノゴケ類	84%
- ANON-2000163-R		
- ANON-2004532-F		
- ANON-2001227-F		

DYW:KP(配列番号:3)

- UWOD-2140463-R	Plagiogyria japonica, シダ植物	86%
------------------	----------------------------	-----

[0081] したがって本発明はまた、下記d～iのいずれか一のポリペプチドを提供する：

d. 配列番号:1の配列と78%を超える配列同一性を有し、好ましくは80%以上、より好ましくは85%以上、さらに好ましくは90%の配列同一性を有し、さらに好ましくは95%の配列同一性を有し、さらに好ましくは97%の配列同一性を有し、かつC-to-U編集活性を有するポリペプチド；

e. 配列番号:2の配列と84%を超える配列同一性を有し、好ましくは85%以上、より好ましくは90%以上、さらに好ましくは95%の配列同一性を有し、さらに好ましくは97%の配列同一性を有し、かつC-to-U編集活性を有するポリペプチド；

f. 配列番号:3の配列と86%を超える配列同一性を有し、好ましくは87%以上、より好ましくは90%以上、さらに好ましくは95%の配列同一性を有し、さらに好ましくは97%の配列同一性を有し、かつU-to-C編集活性を有するポリペプチド。

g. 配列番号:1の配列において、1～29個、好ましくは1～25個、より好ましくは1～21個、さらに好ましくは1～17個、さらに好ましくは1～13個、さらに好ましくは1～9個、さらに好ましくは1～5個のアミノ酸を、のアミノ酸を、

置換、欠失、又は付加した配列を有し、かつC-to-U編集活性を有するポリペプチド;

h. 配列番号:2の配列において、1~21個、好ましくは1~18個、より好ましくは1~15個、さらに好ましくは1~12個、さらに好ましくは1~9個、さらに好ましくは1~6個のアミノ酸を、置換、欠失、又は付加した配列を有し、かつC-to-U編集活性を有するポリペプチド;

i. 配列番号:3の配列において、1~18個、好ましくは1~16個、より好ましくは1~14個、さらに好ましくは1~12個、さらに好ましくは1~10個、さらに好ましくは1~8個のアミノ酸を、さらに好ましくは1~6個のアミノ酸を、置換、欠失、又は付加した配列を有し、かつU-to-C編集活性を有するポリペプチド。

[0082] また、これらの配列のアライメントを下記に示す。アライメントはAliViewを用いて作成した (Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large data sets. *Bioinformatics*30(22): 3276-3278. <http://dx.doi.org/10.1093/bioinformatics/btu531>)。なお、同一のアミノ酸はドットで表した。

[0083]

[0084] [RNA結合ドメイン]

本発明においては、DYWドメインを用いて編集標的を変換する際に、編集標的が含まれるRNAを標的化するために、RNA結合タンパク質が、RNA結合ドメインとして使用される。

[0085] RNA結合ドメインとして用いられるRNA結合タンパク質の好ましい例は、PPRモチーフにより構成されるPPRタンパク質である。

[0086] (PPRモチーフ)

PPRモチーフは、特に記載した場合を除き、Web上のタンパク質ドメイン検索プログラムでアミノ酸配列を解析した際に、PfamにおいてPF01535、Prosit eにおいてPS51375で得られるE値が所定値以下（望ましくはE-03）のアミノ酸配列をもつ30～38アミノ酸で構成されるポリペプチドをいう。本発明で定義するPPRモチーフを構成するアミノ酸の位置番号は、PF01535とほぼ同義である一方で、PS51375のアミノ酸の場所から2引いた数（例；本発明の1番→PS51375の3番）に相当する。ただし、“ii” (-2) 番のアミノ酸というときは、PPRモチーフを構成するアミノ酸の後ろ（C末端側）から2番目のアミノ酸、又は次のPPRモチーフの1番アミノ酸に対して2個N末端側、すなわち-2番目のアミノ酸とする。次のPPRモチーフが明確に同定されない場合、次のヘリックス構造の1番目のアミノ酸に対して、2個前のアミノ酸を“ii”とする。Pfamについては<http://pfam.sanger.ac.uk/>、Prosit eについては<http://www.expasy.org/prosite/>を参照することができる。

[0087] PPRモチーフの保存アミノ酸配列は、アミノ酸レベルでの保存性は低い、2次構造上で2つの α ヘリックスはよく保存されている。典型的なPPRモチーフは35アミノ酸で構成されるが、その長さは30～38アミノ酸と可変的である。

[0088] PPRモチーフは、より具体的には、式1で表される、30～38アミノ酸長のポリペプチドからなる。

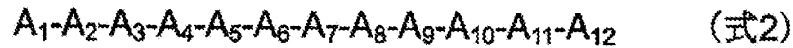
[0089] [化1]



[0090] 式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[0091] [化2]



[0092] 式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[0093] [化3]



[0094] 式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{III} \sim L_{VII}$ は存在しない場合がある。

[0095] (PPR-code)

PPRモチーフは、1、4、ii番の3つのアミノ酸の組み合わせが、塩基との特異的な結合のために重要であり、これらの組み合わせにより、結合する塩基がいずれであるかを決定できる。1、4、ii番の3つのアミノ酸の組み合わせと結合可能な塩基との関係は、PPR-codeとして知られており（前掲特許文献2）、下記のとおりである。

[0096] (1) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン及びアスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にCに、その次にA又はGに対して結合するという、選択的なRNA塩基結合能を有する。

(2) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にGに、その次にCに対して結合するが、Uには結合しないという、選択的なRNA塩基結

合能を有する。

(3) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し、次にA又はUに対して結合するが、Gには結合しないという、選択的なRNA塩基結合能を有する。

(4) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、グルタミン酸、グリシン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合するが、A、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(5) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し、次にUに、その次にAに対して結合するが、Gには結合しないという、選択的なRNA塩基結合能を有する。

(6) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合し、次にUに対して結合するが、AとCには結合しないという、選択的なRNA塩基結合能を有する。

(7) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、リジン、トレオニン、アスパラギン酸、の場合、そのPPRモチーフは、Gに強く結合し、次にAに対して結合するが、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(8) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、セリン、アスパラギン酸の場合、そのPPRモチーフは、Aに強く結合し、次にCに、その次にG及びUに対して結合するという、選択的なRNA塩基結合能を有する。

(9) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、セリンの場合、そのPPRモチーフは、Cに強く結合し、次にUに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(10) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合するが、G、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(11) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にAに対して結合するが、G及びCには結合しないという、選択的なRNA塩基結合能を有する。

(12) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、トレオニン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合するが、G、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(13) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、メチオニン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にCに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(14) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、プロリン、アスパラギン酸の場合PPR、そのモチーフは、Uに強く結合し、次にCに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(15) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、チロシン、プロリン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合するが、A、G及びCには結合しないという、選択的なRNA塩基結合能を有する。

(16) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、ロイシン、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合するが、A、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

[0097] (Pアレイ)

PPRタンパク質は、構成するPPRモチーフの構造によって、PとPLSの2つのファミリーに分類される。

P型のPPRタンパク質は、標準的な35アミノ酸のPPRモチーフ（P）の単純な繰り返し（Pアレイ）から構成されている。本発明のDYWドメインは、P型のPPRタンパク質に連結して用いることができる。

[0098] （PLSアレイ）

PLS型のPPRタンパク質のPPRモチーフの並びは、P1、L1、そしてS1の3つのPPRモチーフの繰り返し単位として構成され、その繰り返しのC末端側に、P2、L2、そしてS2モチーフが続く。さらにこの最後のP2L2S2モチーフのC末端側にE1及びE2と呼ばれる2つのPPR様モチーフ、そしてDYWドメインが続く場合がある。

[0099] P1の全長は、標的塩基に結合することができる限り特に限定されないが、例えば33～37アミノ酸長であり、好ましくは34～36アミノ酸長であり、より好ましくは35アミノ酸長である。L1の全長は、標的塩基に結合することができる限り特に限定されないが、例えば33～37アミノ酸長であり、好ましくは34～36アミノ酸長であり、より好ましくは35アミノ酸長である。S1の全長は、標的塩基に結合することができる限り特に限定されないが、例えば30～33アミノ酸長であり、好ましくは30～32アミノ酸長であり、より好ましくは31アミノ酸長である。

[0100] P2の全長は、標的塩基に結合することができる限り特に限定されないが、例えば33～37アミノ酸長であり、好ましくは34～36アミノ酸長であり、より好ましくは35アミノ酸長である。L2の全長は、標的塩基に結合することができる限り特に限定されないが、例えば34～38アミノ酸長であり、好ましくは35～37アミノ酸長であり、より好ましくは36アミノ酸長である。S2の全長は、標的塩基に結合することができる限り特に限定されないが、例えば30～34アミノ酸長であり、好ましくは31～33アミノ酸長であり、より好ましくは32アミノ酸長である。配列番号:17、22、及び27には、本明細書の実施例の項で使用したP2の配列を示した。配列番号:18、23、及び28には、本明細書の実施例の項で使用したL2の配列を示した。配列番号:19、24、及び29には、本明細書の実施例の項で使用したS2の配列を示した。

- [0101] E1の全長は、標的塩基に結合することができる限り特に限定されないが、例えば32～36アミノ酸長であり、好ましくは33～35アミノ酸長であり、より好ましくは34アミノ酸長である。配列番号:20、25、30には本明細書の実施例の項で使用したE1の配列を示した。
- [0102] E2の全長は、標的塩基に結合することができる限り特に限定されないが、例えば30～34アミノ酸長であり、好ましくは31～33アミノ酸長であり、より好ましくは33アミノ酸長である。配列番号:21、26、31には本明細書の実施例の項で使用したE2の配列を示した。
- [0103] PLS型のPPRタンパク質において、P1L1S1の繰り返し部分とP2までの部分は、上述したPPR-codeのルールに従い、標的RNAの配列に応じて設計することができる。
- [0104] S2モチーフはiiのアミノ酸（配列番号19の31番目のN）が対応するヌクレオチドと相関がある（前掲非特許文献8）。また、S2モチーフの標的塩基の4つ右のC又はUが、DYWドメインによる編集標的塩基であることに留意してPLS型PPRタンパク質に組み込むことができる。
- [0105] E1モチーフでは、ヌクレオチドとの相関は4番目（配列番号:20の4番目のG）のアミノ酸にしか見られない（Ruwe et al. (2019) *New Phytol.* 222 218-229.）。
- [0106] E2モチーフ中の4番目（配列番号:21の4番目のV）及び最後（配列番号:21の33番目のK）のアミノ酸は高度に保存されており、特定のPPR-RNA認識には関与していない（前掲非特許文献2）。
- [0107] P1L1S1の繰り返し数は、標的塩基配列に結合することができる限り特に限定されないが、例えば1～5であり、好ましくは2～4であり、より好ましくは3である。原理的には1単位（3個）でも用い得る。PPRモチーフが5個から構成されるMEF8（L1-S1-P2-L2-S2-E-DYW）は約60箇所の編集に関わることが知られている。
- [0108] 天然のPPRタンパク質では、最初と最後に位置するP1L1S1は、特定の位置にあるアミノ酸残基に明確な違いが見られ、内部のP1L1S1とは異なる。天然に

存在するものになるべく近い人工PLSアレイを設計するとの観点からは、一番目（N末端側）に位置するP1L1S1、内部に位置するP1L1S1、そしてP2L2S2の直前に位置する最後（C末端側）のP1L1S1に分け、PPRモチーフの位置に応じた3種類のP1L1S1を設計するとよい。なお、天然にはP-L-Sの繰り返し単位で構成されているものとは別に、SS(31アミノ酸)が繰り返している場合もあり、このようなものも本発明に利用できる。

[0109] [DYWタンパク質]

本発明は、少なくとも1個のPPRモチーフを含み、PPR-codeのルールにしたがって標的RNAに配列特異的に結合可能なRNA結合ドメイン、及び前述したDYW:PG、DYW:WW、又はDYW:KPのいずれかであるDYWドメインを含む、標的RNAを編集するためのDYWタンパク質を提供する。

[0110] このようなDYWタンパク質は、人工の物でありうる。人工とは、天然物ではなく、人為的に合成したものであることをいう。人工であることは、例えば、天然にはない配列のDYWドメインを有する場合、天然にはない配列のPPR結合ドメインを有する場合、RNA結合ドメインとDYWドメインとの天然にはない組み合わせを有する場合、動物細胞におけるRNA編集用のタンパク質が、植物由来の天然DYWタンパク質には存在しない部分、例えば、核移行シグナル、ヒトミトコンドリア移行シグナルがさらに付加されている場合が当てはまる。核移行シグナル配列としては、例えば、SV40ラージT抗原由来の、PKKKRKV (SEQ ID NO:32)、ヌクレオプラズミンのNLSである KRPAATKKAGQAKKKK (SEQ ID NO:33) などが挙げられる。

[0111] PPRモチーフの数は、標的とするRNAの配列に応じて適宜とすることができる。PPRモチーフの数は、少なくとも1個あればよく、2個以上であってもよい。PPRモチーフが2個でRNAと結合できることが知られている (Nucleic Acids Research, 2012, Vol. 40, No. 6, 2712-2723)。

[0112] DYWタンパク質の好ましい態様の一つは、下記のものである。

少なくとも1個、好ましくは2～25個、より好ましくは5～20個、さらに好ましくは10～18個のPPRモチーフを含み、PPR-codeのルールにしたがって標的RN

Aに配列特異的に結合可能なRNA結合ドメイン、及び前述したDYW:PG、DYW:WW、又はDYW:KPのいずれかであるDYWドメインを含む、標的RNAを編集するための人工DYWタンパク質。

[0113] 他のDYWタンパク質の好ましい態様の一つは、下記のものである。

少なくとも1個、好ましくは2～25個、より好ましくは5～20個、さらに好ましくは10～18個のPPRモチーフを含み、PPR-codeのルールにしたがって動物の標的RNAに配列特異的に結合可能なRNA結合ドメイン（好ましくはPLS型のPPRタンパク質であるRNA結合ドメイン）、及び前述したDYW:PG、DYW:WW、又はDYW:KPのいずれかであるDYWドメインを含む、標的RNAを編集するためのDYWタンパク質。

[0114] 本発明のDYWドメイン、RNA結合ドメインであるPPRタンパク質、及びDYWタンパク質は、当業者にはよく知られた方法で、比較的大量にも調製可能である。そのような方法は、目的のドメイン又はタンパク質が有するアミノ酸配列から、それをコードする核酸配列を決定し、クローニングし、目的のドメイン又はタンパク質を生産する形質転換体を作製することを含みうる。

[0115] [DYWタンパク質の利用]

(DYWタンパク質等をコードする核酸、ベクター、細胞)

本発明は、上述の、PPRモチーフ、DYWタンパク質、又はDYWタンパク質をコードする核酸、核酸を含むベクター（例えば増幅のためのベクター、発現ベクター）も提供する。ベクターには、ウイルスベクターも含まれる。増幅のためのベクターは、大腸菌や酵母を宿主として用いる。本明細書において、発現ベクターとは、例えば上流から、プロモーター配列を有するDNA、所望のタンパク質をコードするDNA、及びターミネーター配列を有するDNAを含むベクターを意味するが、所望の機能を発揮する限り、必ずしもこの順に配列されている必要はない。本発明においては、当業者が通常使用し得る様々なベクターを組み換えて使用することができる。

[0116] 具体的には、本発明は、少なくとも1個のPPRモチーフを含み、PPR-codeのルールにしたがって標的RNA（好ましくは動物の標的RNA）に配列特異的に結

合可能なRNA結合ドメイン（好ましくはPLS型のPPRタンパク質であるRNA結合ドメイン）、及び前述したDYW:PG、DYW:WW、又はDYW:KPのいずれかであるDYWドメインを含む、DYWタンパク質をコードするヌクレオチド配列を提供する。

[0117] また具体的には、本発明は、少なくとも1個のPPRモチーフを含み、PPR-codeのルールにしたがって標的RNA（好ましくは動物の標的RNA）に配列特異的に結合可能なRNA結合ドメイン（好ましくはPLS型のPPRタンパク質であるRNA結合ドメイン）、及び前述したDYW:PG、DYW:WW、又はDYW:KPのいずれかであるDYWドメインを含む、DYWタンパク質をコードするヌクレオチド配列を含む、標的RNAを編集するためのベクターを提供する。

[0118] 本発明のDYWタンパク質は、真核生物（例えば、動物、植物、微生物（酵母、等）、原生生物）の細胞で機能し得る。本発明のDYWタンパク質は、特に、動物細胞内（in vitro又はin vivo）で機能し得る。本発明のDYWタンパク質、又はDYWタンパク質を発現するベクターを導入し得る動物細胞としては、例えば、ヒト、サル、ブタ、ウシ、ウマ、イヌ、ネコ、マウス、ラット由来の細胞を挙げることができる。また、本発明のDYWタンパク質、又はDYWタンパク質を発現するベクターを導入し得る培養細胞としては、例えば、チャイニーズハムスター卵巣（CHO）細胞、COS-1細胞、COS-7細胞、VERO（ATCC CC L-81）細胞、BHK細胞、イヌ腎由来MDCK細胞、ハムスターAV-12-664細胞、HeLa細胞、WI38細胞、HEK293細胞、HEK293T細胞、PER. C6細胞を挙げることができるが、これらに限定されない。

[0119] （用途）

本発明のDYWタンパク質により、標的RNAに含まれる編集標的CをUに、又は編集標的UをCに変換できる。RNA結合性のPPRタンパク質は、オルガネラで見られるすべてのRNA加工のステップ、切断、RNA編集、翻訳、スプライシング、RNA安定化に関与している。

[0120] また本発明のDYWタンパク質により、ミトコンドリアのRNA 1塩基編集ができる。ミトコンドリアは、独自のゲノムを有しており、呼吸やATP生産に関わ

る重要な複合体の構成タンパク質がコードされている。それらの変異により様々な病気が発症することが知られている。本発明を用いた変異修復により、さまざまな病気の処置が期待できる。

[0121] 一方でCRISPR-Casシステムは、Casタンパク質を改変したADARドメインに融合させることで、細胞質におけるC-to-U RNA編集ツールとして開発されている (Abudayyeh et al., 2019年)。しかしながら、CRISPR-Casシステムはタンパク質とガイドRNAで構成されており、ガイドRNAの効率的なミトコンドリア輸送が難しい。一方で、PPRタンパク質は1分子でRNA編集が可能であり、一般的にタンパク質はN末側にミトコンドリア局在シグナル配列を融合することでミトコンドリアへの送達が可能である。そこで、本技術がミトコンドリアのRNA編集に利用できるかどうかを確認するために、MT-ND2とMT-ND5を標的とするPPRをデザインし、PGまたはWWドメインを融合した遺伝子を作成した(図10a)。

[0122] ミトコンドリア標的配列(MTS)とPPR-P配列からなるこれらのタンパク質は、HEK293T細胞に悪影響を与えないように、MT-ND2とMT-ND5の178番目と301番目のコドンの3番目の位置を標的としている(図10a)。HEK293T細胞へプラスミドで導入後、MT-ND2とMT-ND5のmRNA内で編集を確認した(図10b, c)。これらの4つのタンパク質は、ターゲットに対して最高70%の編集活性が検出され、同じmRNA分子に対してオフターゲット変異は検出されなかった(図10bc)。

[0123] したがって、本発明により提供されるRNA塩基の編集方法は、様々な分野で、以下のような利用が期待できる。

[0124] (1)医療

- ・ 特定の疾患に関連した特定のRNAを認識し、編集する。本発明の利用により、1塩基変異による遺伝病を処置しうる。遺伝病における多くの変異の方向はCからUへの変異である。そのため、特にUをCに変換できる本発明の方法は有用でありうる。

[0125] ・ RNAの抑制・発現をコントロールした細胞を作製する。このような細胞には

、分化・未分化状態をモニタリングした幹細胞（例えば、iPS細胞）、化粧品の評価用モデル細胞、創薬のメカニズム解明や薬理試験を目的として、機能性RNAの発現をON/OFFできる細胞が含まれる。

[0126] (2) 農林水産業

- ・農作物、林産物、水産物などにおいて、収量や品質を改善する。
- ・耐病性の向上、環境耐性の向上、向上された又は新たな機能性を有した生物を育種する。

[0127] 例えば、雑種第一代（F1）作物に関し、DYWタンパク質によるミトコンドリアRNAの編集により人工的にF1作物を作出し、収率や品質を改善できる可能性がある。DYWタンパク質によるRNA編集は、従来技術よりも正確かつ迅速に、生物の品種改良、育種（生物を遺伝的に改良すること）が可能である。また、DYWタンパク質によるRNA編集は、遺伝子組み換えのように外来遺伝子により形質を転換するものではないため、変異体の選抜や戻し交雑という旧来の育種の手法に近いといえる。そのため、地球規模の食糧問題、環境問題にも、確実かつ迅速に対応しうる。

[0128] (3) 化学

- ・微生物、培養細胞、植物体、動物体（例えば昆虫体）を利用した有用物質生産において、RNAの操作により、タンパク発現量を制御する。これにより、有用物質の生産性を向上することができる。有用物質の例は、抗体、ワクチン、酵素等のタンパク質性の物質のほか、医薬品の中間体、香料、色素等の比較的低分子の化合物である。

[0129] ・藻類や微生物の代謝経路の改変により、バイオ燃料の産生効率を改善する。

[0130] [用語]

数値範囲x～yは、特に記載した場合を除き、両端の値x及びyを含む。

[0131] タンパク質やポリペプチドのアミノ酸配列に関し、アミノ酸残基を、単に、アミノ酸とすることがある。

[0132] 塩基配列（ヌクレオチド配列ということもある。）又はアミノ酸配列に関

し「同一性」とは、特に記載した場合を除き、2つの配列を最適の態様で整列させた場合に、2つの配列間で共有する一致した塩基又はアミノ酸の個数の百分率を意味する。すなわち、同一性＝（一致した位置の数／位置の全数）×100で算出でき、市販されているアルゴリズムを用いて計算することができる。また、このようなアルゴリズムは、Altschul et al., J.Mol.Biol. 215(1990) 403-410に記載されるNBLAST及びXBLASTプログラム中に組込まれている。より詳細には、塩基配列又はアミノ酸配列の同一性に関する検索・解析は、当業者には周知のアルゴリズム又はプログラム（例えば、BLASTN、BLASTP、BLASTX、ClustalW）により行うことができる。プログラムを用いる場合のパラメーターは、当業者であれば適切に設定することができ、また各プログラムのデフォルトパラメーターを用いてもよい。これらの解析方法の具体的な手法もまた、当業者には周知である。

[0133] 塩基配列又はアミノ酸配列に関し、配列同一性は、特に記載した場合を除き、高い方が好ましい。具体的には、40%以上であることが好ましく、45%以上であることがより好ましく、50%以上であることがさらに好ましく、55%以上であることがさらに好ましく、60%以上であることがさらに好ましく、65%以上であることがさらに好ましい。また、70%以上であることが好ましく、80%以上であることがより好ましく、85%以上であることがさらに好ましく、90%以上であることがさらに好ましく、95%以上であることがさらに好ましく、97.5%以上であることがさらに好ましい。

[0134] ポリペプチド又はタンパク質に関し、「置換、欠失、又は付加した配列」というときの置換等されるアミノ酸の個数は、特に記載した場合を除き、いずれのモチーフ又はタンパク質においても、そのアミノ酸配列からなるモチーフ又はタンパク質が所望の機能を有する限り特に限定されないが、1～9個又は1～4個程度であるか、性質の似たアミノ酸への置換であれば、さらに多くの個数の置換等がありうる。このようなアミノ酸配列に係るポリヌクレオチド又はタンパク質を調製するための手段は、当業者にはよく知られている。

[0135] 性質の似たアミノ酸とは、ハイドロパシー、荷電、pKa、溶解性等の物性が似ているアミノ酸をいい、例えば、次のようなものを指す。

疎水性（非極性）アミノ酸；アラニン、バリン、グリシン、イソロイシン、ロイシン、フェニルアラニン、プロリン、トリプトファン、チロシン

非疎水性アミノ酸；アルギニン、アスパラギン、アスパラギン酸、グルタミン酸、グルタミン、リジン、セリン、トレオニン、システイン、ヒスチジン、メチオニン；

親水性アミノ酸；アルギニン、アスパラギン、アスパラギン酸、グルタミン酸、グルタミン、リジン、セリン、トレオニン；

酸性アミノ酸：アスパラギン酸、グルタミン酸；

塩基性アミノ酸：リジン、アルギニン、ヒスチジン；

中性アミノ酸：アラニン、アスパラギン、システイン、グルタミン、グリシン、イソロイシン、ロイシン、メチオニン、フェニルアラニン、プロリン、セリン、トレオニン、トリプトファン、チロシン、バリン；

含硫アミノ酸：メチオニン、システイン；

含芳香環アミノ酸：チロシン、トリプトファン、フェニルアラニン。

実施例

[0136] [1. DYWドメインの設計と大腸菌内でのRNA編集活性測定]

[結果]

(DYWドメインの設計)

現在利用可能な植物のゲノム情報のほとんどが被子植物のものであるため、人工DYWタンパク質の開発はDYW：PG型に限定される。一方、トランスクリプトームデータは部分配列であるものの、C-to-U及びU-to-C RNA編集の両方を持つ初期の陸上植物種の情報も含まれる。そこで我々は、1000 Plants (1KP) 国際コンソーシアムによって作成されたトランスクリプトームデータセットに基づいて構築されたPPRタンパク質のデータベース（前掲非特許文献11）を用い、完全な人工DYWドメインを設計した。

[0137] 人工DYWドメインを設計するにあたり、シチジンデアミナーゼ様DYWドメイ

ンに加えて、PPR (P2及びL2) モチーフ並びにPPR様 (S2、E1及びE2) モチーフを設計に加えた。これは、これらのモチーフがRNA編集活性に与える重要性が現時点で分かっていないためである。我々は、ツノゴケ類、小葉類及びシダ植物のDYWドメインの系統樹を作成し、DYW : PG、DYW : WW及びDYW : KPの3つのグループに分類した。次いで、これらグループごとのDYWドメインについて新たに系統樹を作成し、人工DYWドメインの設計に利用可能なタンパク質の系統群を選抜した (図1)。DYW : PGグループは配列間の差が大きかったため、設計に用いるタンパク質の系統群を選択するのが困難であった (図1a)。小葉類のタンパク質の配列が多様性に富み、数も多かったため、我々はこのタンパク質に焦点を当てることとした (PG1)。DYW : WWドメインの設計には、ツノゴケ類にのみ見られる分枝が短いタンパク質の系統群を選択した (図1b、WW1)。この系統群は分枝が短いことからタンパク質間の遺伝子の変異が小さいことが推定される。そしてDYW : KPドメインの設計には、シダ植物に特異的なDYWドメインの系統群に着目した。この系統群に含まれるタンパク質の配列は、アミノ酸の変異は大きい、長さは保存されている (図1c、KP1)。

[0138] (RNA結合ドメインの設計)

我々は、各DYWドメインを、植物のゲノム情報から同定されたPPRモチーフ配列 (前掲非特許文献2) に基づいて設計した人工P又はPLSアレイと融合させた (図2a)。PアレイのPPRモチーフは、35アミノ酸のPモチーフのアライメントから得られた共通配列に基づいて構築し、RNA認識を高めるためにいくつかのアミノ酸を置き換えたものである。PLSアレイの設計には、標準的な長さの35アミノ酸のモチーフ (P1及びL1) と31アミノ酸のモチーフ (S1) を選んだ。天然のPPRタンパク質では、最初と最後に位置するP1L1S1は、特定の位置にあるアミノ酸残基に明確な違いが見られ、内部のP1L1S1とは異なる。天然に存在するものになるべく近い人工PLSアレイを設計するために、一番目 (N末端側) に位置するP1L1S1、内部に位置するP1L1S1、そしてP2L2S2の直前に位置する最後 (C末端側) のP1L1S1に分け、PPRモチーフの位置に応じた3種類のP1L1S1を設計した。今後、これらのタンパク質をその構造に基づいて命名す

る。例えば、PアレイにDYW : WWドメインを融合させたものは、P-DYW : WWとする。

[0139] (人工DYWタンパク質は、標的配列を特異的に編集可能である)

シロイヌナズナにおいて、PPRタンパク質CLB19は葉緑体rpoA及びclpP RNA上にあるRNA編集部位を認識する。我々は、rpoA編集部位を標的とするPPRタンパク質を設計することにした。P1、L1、S1及びP2モチーフにおける4番目とii番目のアミノ酸はPPRコードに従った。一方、C末端PPR様モチーフ (L2、S2、E1、E2) に関するPPRコードは分かっていないため、L2、S2、E1及びE2モチーフに関しては、CLB19の4番目とii番目のアミノ酸を使用した (図2b)。

[0140] 組換え人工DYWタンパク質をコードしている遺伝子領域を発現ベクターにクローニングし、その終止コドンの下流に標的配列を付加した。過去に行われたヒメツリガネゴケの2種類のPPRタンパク質 (PPR56、PPR65) に関する研究に基づき (前掲非特許文献12)、設計したPPRタンパク質のRNA編集活性を大腸菌内で試験する方法を開発した。設計したPLS-DYW : PG1及びPLS-DYW : WW1はDNAに対して編集活性が見られなかった一方、RNAに対して90%を上回る編集効率でシチジンをウリジンに置換した (図3a、b、d)。PLSドメインの代わりにPアレイを用いた場合、編集効率が10~40%低下した。一方で、P-DYW : KP及びPLS-DYW : KPの両方でU-to-C RNA編集活性が見られた (図3c、e)。しかしながら、その編集活性はDYW : PG及びDYW : WWと比較して下がっている。

[0141] [方法]

(系統樹)

DYWドメイン (最短132アミノ酸) を含むP2-L2-S2-E1-E2-DYW領域を、PPRデータベース (<https://ppr.plantenergy.uwa.edu.au/onekp/>; 前掲非特許文献11) から抽出した。得られた配列のアライメントをMAFFT L-INS-i (v7.407 automatic mode) (K. Katoh, D. M. Standley (2013). Mol Biol Evol. 30(4): 772-780.) を使って作成し、その後、trimAl (v.1.4.rev15) (Salvador Capella-Gutierrez, et al. (2009). Bioinformatics. 25(15): 1972-1973.) を用いてトリミングした。トリミングではパラメーターをgt 0.2 cons

20とした。活性部位 (HxEx_nCxxCH) に変異を持つ配列はアライメントから除いた (図4)。

[0142] 残りの配列から、FastTree (v2.1.10) (Price, M. N., et al. (2010). PLoS One 5:e9490.) を使用して系統樹を作成し、DYW : PG、DYW : WW及びDYW : KPを同定した。このときのパラメーターはwag及びcat 8とした。

[0143] (Trx-PPR-DYWタンパク質及び標的配列のクローニング)

各DYWドメイン (P2、L2、S2、E1、E2を含む) の共通配列をEMBOSS : cons (v.6.6.0.0) を使って設計した。タンパク質発現用ベクターはpET21b+PAを改変して、元のEsp3IとBpiI制限酵素部位を除去し、クローニングサイトとして2つのEsp3I部位を付加した。遺伝子は4つのセクション (Trx、PPRアレイ、DYWドメイン及びRNA編集部位) に分けて、2段階のGolden Gate法を使って構築した。初めに、改変pET21bベクターのEsp3I部位に、1) チオレドキシン-6×His-TEV遺伝子領域 (3' にBpiI制限酵素部位を含む)、2) P2-L2-S2-E1-E2-DYW遺伝子領域 (5' にBpiI制限酵素部位を含む)、及び3) RNA編集部位のコード配列領域、の3部分をクローニングした。次いで、BpiI部位に完全長PLSドメイン又はPドメインをクローニングし、PLS : DYW (配列番号:35~37) 又はP : DYW (配列番号:38~40) タンパク質を作製した。

[0144] (大腸菌内でのRNA編集活性測定)

組換えタンパク質のRNA編集活性を大腸菌内で分析するために、我々は、Odenkottらによって開発されたプロトコル (前掲非特許文献12) を改変した。上記で作成したプラスミドDNAを大腸菌Rosetta 2株に導入し、1mLのLB培地 (カルベニシリン50 µg/mL及びクロラムフェニコール17 µg/mL含有) で37 °C、一晚培養した。5mLの適切な抗生物質を含むLB培地を深底24ウェルプレートに準備し、前培養液100 µLをここに播種した。この培養液を、吸光度 (OD₆₀₀) が0.4~0.6に達するまで、37 °C、200rpmで生育させた後、プレートを4 °Cで10~15分間冷却した。次いで、0.4mMのZnSO₄と0.4mMのIPTGを加え、さらに16 °C、180rpmで18時間培養した。培養液750 µLを採取して遠心分離した後、菌体ペレットを液体窒素中で凍結させ、-80 °Cで保存した。

[0145] 凍結した細胞塊を200 μ Lの1-チオグリセロール／ホモジナイズ溶液に再懸濁し、40Wで10秒間、超音波処理して細胞を分離した後、200 μ Lの溶解バッファーを加えた。Maxwell（登録商標） RSC simplyRNA Tissue Kit（プロメガ）を使いRNAを抽出した。RNAをDNaseI（タカラバイオ）で処理し、SuperScript（登録商標） III Reverse Transcriptase（インビトロジェン）を使い、処理したRNA 1 μ gと1.25 μ Mのランダムプライマー（6 mer）を用いてcDNAを合成した。NEBNext High-Fidelity 2x PCR Master Mix（ニューイングランドバイオラボ）、1 μ LのcDNA及び、チオレドキシンとT7ターミネーター配列のプライマーを使って編集部位を含む領域を増幅した。NucleoSpin（登録商標） Gel and PCR Cleanup（タカラバイオ）でPCR産物を精製し、DYWドメイン配列に特異的なフォワードプライマーを用いてシーケンス解析を行い、RNA編集部位の塩基を決定した。RNA編集効率の測定には、編集部位のCとUの波形ピークの高さの比を用い、C-to-U RNA編集効率は $U/(C+U) \times 100$ 、U-to-C編集効率は $C/(C+U) \times 100$ によって算出した。独立した実験を3回繰り返した。

[0146] [引用した配列の一覧]

[0147]

[表3-1]

PLS array (3 P1L1S1 triplets)

SEQ ID NO:

P1	FSWNSMIRGYARSGQPPEEALSLFSQMRRSGIKPDS	4
L1	YTFPFVLKACASLSSLEEGKQIHAVIKSGFESDV	5
S1	YVQSSLIDMYAKCGSLEDARKVFDEMPERNV	6
P1	VSWNAMISGYAQNGQSEEALSELFREMQKEGIKPSE	7
L1	FTFCSVLKACASLGSLEMKGQIHGYVIKSGFESIV	8
S1	FVGNALIDMYAKCGSLEDARKVFDEMPERTV	9
P1	VSWTAMISGYAQNGQSEEALSELFREMQREGVKPDE	10
L1	VTLPVLSACANLGALEQGGKQIHAYVIKNGFESDV	11
S1	FVGSALIDMYAKCGSIEEARKVFDEMPEKDV	12

P array (10 PPR motifs)

SEQ ID NO:

P	VTYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNV	13
P	VTYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDV	14
P	VTYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDV	14
P	VTYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNV	13
P	VTYNTLIDGLCKSGKIEEALKLFKEMEKGITPSV	15
P	VTYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNV	13
P	VTYNTLIDGLCKSGKIEEALKLFKEMEKGITPSV	15
P	VTYTTLIDGLCKAGKVDEALELFDKMERGIKPDV	16
P	VTYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDV	14
P	VTYTTLIDGLCKAGKVDEALELFDKMERGIKPDV	16

[0148]

[表3-2]

DYW:PG		SEQ ID NO:
P2	VSWNAMIAAYAQHGKGKEALQLFQQMQQEGVKPSE	17
L2	VTFTSILSACSHAGLVDEGHYFESMSPDYGITPRV	18
S2	EHYGCMVDLLGRAGRLDEAEDLIKSMFPQPNV	19
E1	VVWGTLLGACRVHGDVERGERAAERILELDPESA	20
E2	APYVLLSNIYAAAGRWDEAAKVRKLMKERG VKK	21
DYW	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLS KQMKKEAGYVPD TKFVLHDVEEEEKEQLLCYHSEKLAI AFGLISTPPGTPLRIIKNLRVCGDCHTATKFISKIVG REIVVRDANRFHHFKDGVCSGDIW	1
DYW:WW		SEQ ID NO:
P2	VTWNALIAGYARQGESDLVFHLLERMQRQEGIQPSG	22
L2	VTFTSVLTVCSHAGLVDEGQKYFDAMSEDYGITPRI	23
S2	EHYGCMVDLLGRAGQMDEAVAMVEKMPFPNL	24
E1	VTWGTLLGACRKWNNVEIGRHAFECAVRLDEKSA	25
E2	AAAYVLSNIYADAHMWEERDKIQAMRK NARAWK	26
DYW	KPGQSWWTD TG GIVHTFVVGDTKHPQSQDIYAKLKDL YVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHSEKLA IACALINTPPGTPIRIVKNLRVCDDCHKAIALISKIE GRNIICRDASRFHNYKDGKCSCGDIW	2
DYW:KP		SEQ ID NO:
P2	VAWNSMIKGYGMNHEGKMAVQC FEDMQKQGVKPSA	27
L2	VTFTCLLTACSHASLVSEGQ EYFKMMREEYGIAPRV	28
S2	EHYGCMVDLLARS GHLYEAEKFLEMLCPPNE	29
E1	GTWGALLSACKTYGEVELGLRCFQQLVQLNPESA	30
E2	AWYVLMADIYAGAGRWDDAYRIEELRKHAGAKK	31
DYW	KPASALIEVNKKVHEFVVGNNQSEEISAMLKSLNSRM KEEGHVPNL DLVLKPVSD EEKEAALCEHA EKLAIAFG LLNTPQGQTLRVTKNLRMCNDCHNASKIISKIERREI ILRDDCCIHHFKDGLCSCGDMF	3

[表3-3]

<PLS array (3 P1L1S1 triplets) + DYW:PG-lyc> VFSWNSMIRGYARNQPEEALSLYSQMRRSGIKPDNYTFPFVLKACASLSSLKEGKQIHGHVKS GFESDVYVQSALIDMYAKCGE LEDARKVFDEMPERNVVSWNAMISGYAQNQGSEAELELFREMQQEGIKPSEFTFCSVLSACASLGSLEMCKQIHGYVKS GFESIV FVGNALIDMYAKCGSIEDARKVFDEMPRTVVSWTAMISGYAQNQGSEAELELFREMQRGVKPDVTLPSVLSACANLGALEQKG QIHAYVKS GFESDVYVQSALIDMYAKCGSIEDARKVFDKMPERDVVSWNAMIAAYAQHGHGKEALQLFQQMQQEGVKKPSEVTFTS ILSACSHAGLVDEGHYFESMSPDYGITPRVEHYGCMVDLLGRAGRLDEAEDLIKSMFPQPNVYVWGTLGACRVHGDVERGERAA ERILELDPESAAPYVLLSNIIYAAAGRWDEAAKVRKLMKRGVKKPGCSWIEVNNKVHEFVAGDKSHPQTKIYAELERLSKQMK AGYVPDTKFLVLDHVEEEKEQLLCYHSEKLAIAFGLISTPPGTPLRIKLNLRVCGDCHTATKFTISKIVGREIVVRDANRFHHFKDG VCSGCDYW (SEQ ID NO:35)
<PLS array (3 P1L1S1 triplets) + DYW:WW-WW5> VFSWNSMIRGYARNQPEEALSLYSQMRRSGIKPDNYTFPFVLKACASLSSLKEGKQIHGHVKS GFESDVYVQSALIDMYAKCGE LEDARKVFDEMPERNVVSWNAMISGYAQNQGSEAELELFREMQQEGIKPSEFTFCSVLSACASLGSLEMCKQIHGYVKS GFESIV FVGNALIDMYAKCGSIEDARKVFDEMPRTVVSWTAMISGYAQNQGSEAELELFREMQRGVKPDVTLPSVLSACANLGALEQKG QIHAYVKS GFESDVYVQSALIDMYAKCGSIEDARKVFDKMPERDVVSWNAMIAAYAQHGHGKEALQLFQQMQQEGVKKPSEVTFTS VLTVC SHAGLVDEGQRYFDAMSEYDITPRIEHYGCMVDLLGRAGQMDAAMVEKMPFQPNLVTWGTLGACRVHGDVERGERAA ECARVLDKSAAYVLMNSIYADAHMWEERDKIQAMRKNARAWKKPGQSWWTDTGGLVHTFVVGDTKHPQSQDIYAKLKDLYVKMK EEGYVPHLDCVLWDISDDEKEDALCGHSEKLAIAFGLINTPPGTPIRIVKNLRVCGDCHKAIALISKIEGRNIIICRDASRFHNYKD GKCSGCDYW (SEQ ID NO:36)
<PLS array (3 P1L1S1 triplets) + DYW:KP-lepto> VFSWNSMIRGYARNQPEEALSLYSQMRRSGIKPDNYTFPFVLKACASLSSLKEGKQIHGHVKS GFESDVYVQSALIDMYAKCGE LEDARKVFDEMPERNVVSWNAMISGYAQNQGSEAELELFREMQQEGIKPSEFTFCSVLSACASLGSLEMCKQIHGYVKS GFESIV FVGNALIDMYAKCGSIEDARKVFDEMPRTVVSWTAMISGYAQNQGSEAELELFREMQRGVKPDVTLPSVLSACANLGALEQKG QIHAYVKS GFESDVYVQSALIDMYAKCGSIEDARKVFDKMPERDVVSWNAMIAAYAQHGHGKEALQLFQQMQQEGVKKPSEVTFTS LLTACSHASLVSEGQYFKMMREEYGIAPRVEHYGCMVDLLARSCHLYEAEKFLMLCPNNEGTVGALLSACKTYGEVELGLRCFQ QLVQLNPESAAYVLMADIYAGAGRWDDAYRIEELRKHAGAKKPPASALIEVNNKVHEFVVGNNQSEIISAMLSLNSRMKEEGHV PNLDLVLKPVSDKEEAALCEHAELATAFGLINTPPGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHFKDGLCSG GDMF (SEQ ID NO:37)
<P array (10 PPR motifs) + DYW:PG-lyc> VVVYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGR LDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEKGK TPSVVYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEKGKITPSVVYTTLIDGLCK AGKVDEALELFDEMKEGKIPDVVYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGKVDEALELFDEMKE RGIKPDVVSWNAMIAAYAQHGHGKEALQLFQQMQQEGVKKPSEVTFTSILSACSHAGLVDEGHYFESMSPDYGITPRVEHYGCMVD LLGRAGRLDEAEDLIKSMFPQPNVYVWGTLGACRVHGDVERGERAAERILELDPESAAPYVLLSNIIYAAAGRWDEAAKVRKLMKE RGVKKPGQCSWIEVNNKVHEFVAGDKSHPQTKIYAELERLSKQMKAGYVPDTKFLVLDHVEEEKEQLLCYHSEKLAIAFGLIST PPGTPLRIKLNLRVCGDCHTATKFTISKIVGREIVVRDANRFHHFKDGVCSGCDYW (SEQ ID NO:38)
<P array (10 PPR motifs) + DYW:WW-WW5> VVVYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGR LDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEKGK TPSVVYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEKGKITPSVVYTTLIDGLCK AGKVDEALELFDEMKEGKIPDVVYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGKVDEALELFDEMKE RGIKPDVVTWNALIAGYARQGESDLVFLHLLMRQEGIQPSGVTFTSVLTVC SHAGLVDEGQRYFDAMSEYDITPRIEHYGCMVD LLGRAGQMDAAMVEKMPFQPNLVTWGTLGACRVHGDVERGERAAERILELDPESAAYVLMNSIYADAHMWEERDKIQAMRKN ARAWKKPGQSWWTDTGGLVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDISDDEKEDALCGHSEKLAIAFGLINT PPGTPIRIVKNLRVCGDCHKAIALISKIEGRNIIICRDASRFHNYKDGKCSGCDYW (SEQ ID NO:39)
<P array (10 PPR motifs) + DYW:KP-lepto> VVVYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGR LDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEKGK TPSVVYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEKGKITPSVVYTTLIDGLCK AGKVDEALELFDEMKEGKIPDVVYNTLIDGLCKAGRLDEAEELLEEMEKGKIPDVVYNTLIDGLCKAGKVDEALELFDEMKE RGIKPDVVAWNSMIGYGMNHEGKMAVQCQFEDMQQGVKPSAVTFTCLLTACSHASLVSEGQYFKMMREEYGIAPRVEHYGCMVD LLARSCHLYEAEKFLMLCPNNEGTVGALLSACKTYGEVELGLRCFQQLVQLNPESAAYVLMADIYAGAGRWDDAYRIEELRKHA GAKKKPPASALIEVNNKVHEFVVGNNQSEIISAMLSLNSRMKEEGHVPNLDLVLKPVSDKEEAALCEHAELATAFGLINTPPGQ TLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHFKDGLCSGDMF (SEQ ID NO:40)

[0150] [2. 動物培養細胞での実施例]

[結果]

PLS型PPRと各DYWドメイン（PG1（配列番号:1）、WW1（配列番号:2）、KP1（配列番号:3））を融合した遺伝子と、標的配列を含むプラスミドをHEK293T

細胞にトランスフェクションし、培養後RNAを回収した。標的部位のシチジン(C)からウリジン(U)、またはウリジン(U)からシチジン(C)への変換効率をサンガーシーケンシングにより解析した(図5)。PG1またはWW1ドメインを融合した場合、90%以上のC to U活性があり、一方でU to C活性は検出されなかった(図5a, c)。KP1ドメインを融合した場合、U to C活性が25%検出されたが、C to U活性も10%程度検出された(図5b, c)。これらのことから、動物培養細胞においても編集酵素が機能することが分かった。

[0151] [方法]

(動物培養細胞試験用PPR発現プラスミドの作成)

図3で使用したプラスミドから、6xHis-PPR-DYWタンパク質(大腸菌での実験に用いたものと同じタンパク質(配列番号:35~37))遺伝子配列と編集サイトを含む領域(配列番号:34)をPCRで増幅し、Golden Gate Assembly法により動物細胞発現ベクターへクローニングした。ベクターは、CMVプロモーターとヒトβ-グロビンキメライントロンを含むプロモーター制御化でPPRが発現し、SV40 polyadenylation signal によってpoly Aシグナルが付与される。

[0152] (HEK293T細胞培養)

HEK293T細胞は、高グルコース、グルタミン、フェノール-RED、ピルビン酸ナトリウム(富士フイルム和光純薬株式会社)を含むDulbecco's Modified Eagle Medium (DMEM)培地に10%牛胎児血清(カプリコム)と1% ペニシリン-ストレプトマイシン(富士フイルム和光純薬株式会社)を加えたものを用い、37℃、5% CO₂で培養した。細胞は80-90%コンフルエントになった時点で2-3日ごとに継代した。

[0153] (トランスフェクション)

RNA編集アッセイでは、24ウェル平底細胞培養プレートの各ウェルにHEK293T細胞を約 8.0×10^4 個入れ、37℃、5% CO₂で24時間培養した。各ウェルに500 ngのプラスミドを18.5 μlのOpti-MEM(登録商標) I Reduced Serum Medium(ThermoFisher)と1.5 μlのFuGENE(登録商標) HD Transfection Reagent(Promega)を加え、最終的に25 μlの容量とした。混合液は、細胞に添加する

前に、室温で10分間インキュベートした。トランスフェクションの24時間後に細胞を回収した。

[0154] (RNA抽出、逆転写、シーケンシング)

上述の大腸菌でのアッセイと同様に行った。なお、以下の実施例では、特に記載した場合を除き、大腸菌でのアッセイおよび本実施例と同様の方法で実験を行った。

[0155] [3. ドメインスワップによるRNA編集活性への影響]

DYWドメインは、アミノ酸配列の保存性からいくつかの領域に区分けすることができるが、これらのRNA編集活性との関連性は分かっていない。ここでは、KP1 とWW1ドメインの一部をスワップし、HEK239T細胞中で、RNA編集活性への影響を調べた(図6)。WW1ドメインのPG boxとDYWをKP1ドメインのActive siteを含む中央部位に融合したもの(chimKP1a 配列番号:90)において、KP1ドメインよりも高い50% 近いU to C活性が認められ、さらにC to U活性がほとんどなくなることが分かった。ドメインスワッピングにより、KPドメインのU to C編集の性能を改善することに成功した。

[0156] [4. 変異導入によるKPドメインのRNA編集活性の改善]

KPドメインに様々な変異を導入し、KPドメインのRNA編集活性の向上を目指した。KP2～KP23(配列番号:68～89)をデザインし、C to U または U to C のRNA編集活性を、大腸菌(図7a)及びHEK293T細胞(図7b、c)で調べた。KP22(配列番号:88)が最もU to Cへの編集活性が高く、C to Uへの編集活性が低く、KP1と比較してRNA編集活性を改善することに成功した。

[0157] [5. 変異導入によるPGドメインのRNA編集活性の改善]

PGドメインに様々な変異を導入し、PGドメインのRNA編集活性の向上を目指した。PG2～PG13(配列番号:41～53)をデザインし、C to U のRNA編集活性を大腸菌(図8a)及びHEK293T細胞(図8b、c)で調べた。PG11(配列番号:50)が最もC to Uへの編集活性が高く、RNA編集活性を改善することに成功した。

[0158] [6. 変異導入によるWWドメインのRNA編集活性の改善]

WWドメインに様々な変異を導入し、WWドメインのRNA編集活性の向上を目指した。WW2～WW14をデザインし、C to U のRNA編集活性を大腸菌（図9a）及びHEK293T細胞（図9b、c）で調べた。WW11（配列番号:63）が最もC to Uへの編集活性が高く、WW1と比較してRNA編集活性を改善することに成功した。

[0159] [7. PPRタンパク質を用いたヒトミトコンドリアRNA編集]

[結果]

ミトコンドリアは、独自のゲノムを有しており、呼吸やATP生産に関わる重要な複合体の構成タンパク質がコードされている。それらの変異により様々な病気が発症することが知られており、変異修復方法が求められている。

[0160] CRISPR-Casシステムは、Casタンパク質に改変したADARドメインに融合させることで、細胞質におけるC-to-U RNA編集ツールとして開発されている（Abu dayyeh et al. 2019 Science Vol.365, Issue 6451, pp.382-386）。しかしながら、CRISPR-Casシステムはタンパク質とガイドRNAで構成されており、ガイドRNAの効率的なミトコンドリア輸送が難しい。一方で、PPRタンパク質は1分子でRNA編集が可能であり、一般的にタンパク質はN末側にミトコンドリア局在シグナル配列を融合することでミトコンドリアへの送達が可能である。そこで、本技術がミトコンドリアのRNA編集に利用できるかどうかを確認するために、MT-ND2とMT-ND5を標的するPPRをデザインし、PG1またはWW1ドメインを融合した遺伝子を作成した（図10a）。

[0161] ミトコンドリア標的配列(MTS)とP-DYW配列からなるこれらのタンパク質は、HEK293T細胞に悪影響を与えないように、MT-ND2とMT-ND5の178番目と301番目のコドンの3番目の位置を標的としている（図10a）。HEK293T細胞へプラスミドで導入後、MT-ND2とMT-ND5 mRNA内で編集を確認した（図10b, c）。これらの4つのタンパク質は、ターゲットに対して最高70%の編集活性が検出され、同じmRNA分子に対してオフターゲット変異は検出されなかった（図10bc）。

[0162] [方法]

（ミトコンドリア編集のためのクローニング）

Zea maysのLOC100282174タンパク質からのミトコンドリア標的配列 (Chin et al. 2018)、10個のPPR-P及びPPR様モチーフ、及びDYWドメイン部分 (DYWドメインはPG1とWW1を使用) を、Golden Gate AssemblyによりCMVプロモーターの制御下の発現プラスミドにクローニングした (配列番号:91~94)。

[0163] [引用した配列の一覧]

[0164] [表4-1]

	Sequence	SEQ ID NO:
PG1	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	1
PG2	QPGCSSIEVDNKNVHEFVAGDRSHPQTKEIYAELERLNKMKEAGYVPDTRLVLHDVDEEKEHLLCHHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKYISKIVGREIVVRDANRFHHFKDGVCSGDIY	41
PG3	KPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	42
PG4	EPACSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	43
PG5	EPGCWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	44
PG6	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	45
PG7	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLNKMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	46
PG8	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	47
PG9	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHPEVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	48
PG10	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRVTKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	49
PG11	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRMCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	50
PG12	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANRFHHFKDGVCSGDIY	51
PG13	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDDCRFHHFKDGVCSGDIY	52
PG14	EPGCSWIEVNNKVHEFVAGDKSHPQTKEIYAELERLSKQMKEAGYVPDTKFVLHDVEEEEKEQLLCYHSE KLAIAFGLISTPPGTPLRI IKNLRVCGDCHTATKFI SKIVGREIVVRDANCVHHFKDGVCSGDIY	53

[0165]

[表4-2]

	Sequence	SEQ ID NO:
WW1	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	2
WW2	QPGQSWWTD MGGVVHTFVVGDRHPESQDIYAKLKDLYVKMKEEGYVPHLDSVLRDIPDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	54
WW3	EPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	55
WW4	KPAQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	56
WW5	KPGQAWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	57
WW6	KPGQSTWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	58
WW7	KPGQSWTDTGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHSE KLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	59
WW8	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	60
WW9	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	61
WW10	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWPI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	62
WW11	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	63
WW12	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	64
WW13	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	65
WW14	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	66
WW15	KPGQSWWTD TGGIVHTFVVGDTKHPQSQDIYAKLKDLYVKMKEEGYVPHLDCVLWDI SDDEKEDALCGHS EKLAIAACALINTPPGTPIRIVKNLRVDDCHKAIALISKIEGRNII CRDASRFHNYKDGKCSGDIY	67

[0166]

[表4-3]

	Sequence	SEQ ID NO:
KP1	KPASALIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	3
KP2	KPAVAGIEVDRHVHTFVVGEERDISCKVRSNMTRLKEEGHVPHTLVKPLSEKEKEDALCGHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	68
KP3	KPAVATIEVNKKVHRFLVGERREDIALKVESTNARLKQEGGHVPHTLVKPLSEKEKEDALCGHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	69
KP4	KPAVATIEVNKKVHRFLVGERREDIALKVESTNARLKQEGGHVPHTLVKPLSEKEKEDALCGHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	70
KP5	KPGSALIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	71
KP6	KPAQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	72
KP7	KPASSLIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	73
KP8	KPGQAWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	74
KP9	KPASALIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	75
KP10	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	76
KP11	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	77
KP12	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	78
KP13	KPASALIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	79
KP14	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	80
KP15	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	81
KP16	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	82
KP17	KPASALIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	83
KP18	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	84
KP19	KPASALIEVNKKVHEFVVGNQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDASCIHHFKDGLCSCGDMF	85
KP20	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDASCIHHFKDGLCSCGDMF	86
KP21	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	87
KP22	KPGQSWTDTGGIVHTFVVGDTKHPQSQDIYAKLSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	88
KP23	KPGQSWTDTGGIVHTFVVGDTKHPQSQDIYAKLSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	89
chim	KPGQSWTDTGGIVHTFVVGDTKHPQSEIEISAMKSLNSRMKEEGHVPNLDLVLKPVSDDEEKEAALCEHA	
KP1a	EKLAI AFGLLNT PQGQTLRVTKNLRMCNDCHNASKIISKIERREIILRDDCCIHHFKDGLCSCGDMF	90

[0167]

[表4-4]

	Sequence	SEQ ID NO:
MTS-P-PGI_ND 2	MALLRAAVSELRRRGRGALTPLPALSSLLSSLSRSPASTRPEPNNPHADRRHVIALRRCPPPLPASAVLAP ELLHARGLLPRHWSHASPLSTSSSSSRPADKAQLTWVDKWIPEAARPYSHGSGSGGGLHMGSGGLSSGDY KDHDGDYKDHDIDYKDDDDKSGSGSGSGSGNSVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYN TLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYTT LIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGKVDEALELFDEMKEKGIKPDVVTYNTL IDGLCKSGKIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYTTLI DGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYTTLI GLCKAGKVDEALELFDEMKEKGIKPDVSVWNAMIAAYAQHGHGKEALQLFQQMQQEGVKPSEVTTPTSILSA CSHAGLVDEGHHYFESMSPDYGITPRVEHYGCMVDLLGRAGRLDEAEDLIKSMFPQPNVVVWGTLGACRV HGDVERGERAAERILELDPESAAPYVLLSNIIYAAAGRWDEAAKVRKLMKERGVKKPEPGCSWIEVNNKVHEF VAGDKSHFQTKETIYAELERLSKQMKAGYVPDTKFLVHDVEEEKEQLLCYHSEKLAIAFGLISTPPGTPL RIKKNLRVCGDCHTATKFIKIVGREIVVRDANRFHFKDGVCSGCDYW	93
MTS-P-WW1_ND 2	MALLRAAVSELRRRGRGALTPLPALSSLLSSLSRSPASTRPEPNNPHADRRHVIALRRCPPPLPASAVLAP ELLHARGLLPRHWSHASPLSTSSSSSRPADKAQLTWVDKWIPEAARPYSHGSGSGGGLHMGSGGLSSGDY KDHDGDYKDHDIDYKDDDDKSGSGSGSGSGNSVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYN TLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYTT LIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGKVDEALELFDEMKEKGIKPDVVTYNTL IDGLCKSGKIEEALKLFKEMEEKGITPSVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYTTLI DGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYTTLI GLCKAGKVDEALELFDEMKEKGIKPDVVTWNALIAIYARQGESDLVFHLLERMRQEGIQPSGVTTPTSILTV CSHAGLVDEGQKYFDAMSEDYGITPRIEHYGCMVDLLGRAGQMDAEMVEMKMPQPNLVTVGTLGACRK WNNVEIGRHAFECARLDEKSAAYVLMNSIYADAHMWEERDKIQAMRKNARAWKKPGQSWWTDGGSIVHT FVVGDTKHPQSQDIYAKLDLYVKMKEGYVPHLDCVLWDISDDEKEDALCGHSEKLAIAACALINTPPGTPL IRIVKNLRVCGDDCHKAIALISKIEGRNIICRDASRFHNYKDGKCSGCDYW	94
MTS-P-PGI_ND 5	MALLRAAVSELRRRGRGALTPLPALSSLLSSLSRSPASTRPEPNNPHADRRHVIALRRCPPPLPASAVLAP ELLHARGLLPRHWSHASPLSTSSSSSRPADKAQLTWVDKWIPEAARPYSHGSGSGGGLHMGSGGLSSGDY KDHDGDYKDHDIDYKDDDDKSGSGSGSGSGNSVVTYNTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYN TLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYNTLIDGLCKAGKVDEALELFDEMKEKGIKPDVVTYTT LIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYTT IDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYNTLI DGLCKSGKIEEALKLFKEMEEKGITPSVVTYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYTTLI GLCKAGKVDEALELFKEMRSKGVKPNVVSWTAMIAAYAQHGHGKEALQLFQQMQQEGVKPNEVTTPTSILSA CSHAGLVDEGHHYFESMSPDYGITPRVEHYGCMVDLLGRAGRLDEAEDLIKSMFPQPNVVVWGTLGACRV HGDVERGERAAERILELDPESAAPYVLLSNIIYAAAGRWDEAAKVRKLMKERGVKKPEPGCSWIEVNNKVHEF VAGDKSHFQTKETIYAELERLSKQMKAGYVPDTKFLVHDVEEEKEQLLCYHSEKLAIAFGLISTPPGTPL RIKKNLRVCGDCHTATKFIKIVGREIVVRDANRFHFKDGVCSGCDYW	95
MTS-P-WW1_ND 5	MALLRAAVSELRRRGRGALTPLPALSSLLSSLSRSPASTRPEPNNPHADRRHVIALRRCPPPLPASAVLAP ELLHARGLLPRHWSHASPLSTSSSSSRPADKAQLTWVDKWIPEAARPYSHGSGSGGGLHMGSGGLSSGDY KDHDGDYKDHDIDYKDDDDKSGSGSGSGSGNSVVTYNTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYN TLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYNTLIDGLCKAGKVDEALELFDEMKEKGIKPDVVTYTT LIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKSGKIEEALKLFKEMEEKGITPSVVTYTT IDGLCKAGKVDEALELFKEMRSKGVKPNVVTYNTLIDGLCKAGRLDEAEELLEEMEEKGIKPDVVTYNTLI DGLCKSGKIEEALKLFKEMEEKGITPSVVTYTTLIDGLCKAGKVDEALELFKEMRSKGVKPNVVTYTTLI GLCKAGKVDEALELFKEMRSKGVKPNVVTWTALIAIYARQGESDLVFHLLERMRQEGIQPNVTTPTSILTV CSHAGLVDEGQKYFDAMSEDYGITPRIEHYGCMVDLLGRAGQMDAEMVEMKMPQPNLVTVGTLGACRK WNNVEIGRHAFECARLDEKSAAYVLMNSIYADAHMWEERDKIQAMRKNARAWKKPGQSWWTDGGSIVHT FVVGDTKHPQSQDIYAKLDLYVKMKEGYVPHLDCVLWDISDDEKEDALCGHSEKLAIAACALINTPPGTPL IRIVKNLRVCGDDCHKAIALISKIEGRNIICRDASRFHNYKDGKCSGCDYW	96